前沿探索: 使用深度强化学习进行蒙托卡洛自适应光场采样与重建

霍宇驰
浙江大学

1 自适应采样与重建

本文旨在探索总结使用深度强化学习进行蒙托卡洛自适应光场采样与重建方向的前沿进展。在使用蒙托卡洛算法进行全局光照明绘制时，如果路径跟踪产生的样本点不足时，绘制结果会包含有大量的噪点，严重影响结果的可用性。为了这一问题，常见的解决方法有两种：一是通过神经网络或者其他方法对绘制结果进行降噪滤波，这可以看作是一种后处理的重建方案 [12]；本文介绍第二种方法，即在采样过程中对路径追踪生成蒙托卡洛样本的过程加以引导，以提高最终绘制结果的质量，这是对采样进行过程进行优化的方案。

绘制技术是图像学领域的核心问题之一，所要解决的问题是如何高效高质地计算光能在三维世界中的传播，以使用更短的时间生成具有物理真实感的图片。但是进行物理真实感的绘制是非常耗时的，需要使用蒙托卡洛积分来求解著名的渲染方程 [14]，这一积分通常需要大量的采样样本才能达到收敛。已经对各种方法进行了深入研究处理这个问题。在需要交互帧率的实时绘制应用领域，对含有噪点的绘制结果进行有偏的图像空间滤波（用于显示单张图片）和光场空间滤波（用于显示光场）是常见的解决方案。然而，有偏绘制不适用于设计、物理仿真、深度学习的训练数据生成、以及其他对质量有较高要求绘制应用中。在某些绘制应用，数学偏差、系统误差或者微小方差造成时域不连续现象都是不可接受的。此时，可以先重建像素上的入射辐射场（光场），和物体表面的双向反射分布函数（BRDF）进行混合，作为引导蒙托卡洛采样的概率分布函数（PDF）。如果重建的 PDF 能够很好地逼近真实的绘制函数分布，就能通过采样很少的蒙托卡洛样本生成数学上无偏的绘制结果。

学界一直在关注光场自适应采样和重建方法。Jensen 等人提出用光子追踪算法估计入射光场 [13]，通过计算入射球面上的光子数直方图，生成能够引导路径追踪采样方向的概率函数。基于这一工作，Budge 等人提出了一种特殊的 PDF 表示来有效地渲染焦散效果 [2]。Vorba 等人提出了使用在线训练方法迭代更新高斯混合模型 (GMM) 的参数，以估计入射光场 PDF[19]。后续的工作基于俄罗斯轮盘赌算法改进了这种方法，并将之扩展到双向散射分布函数 [6, 20]。Hua 等人使用虚拟点光源代替光子追踪算法对入射辐射度进行逼近 [7]。最近，Müller 等人提出了使用空间方向四维空间中的混合四叉树数据结构来表示可在线学习的入射光场 [17]。此外，在线的学习回归方法也被用于采样面向直接照明的入射光场函数以引导采样 [18]。

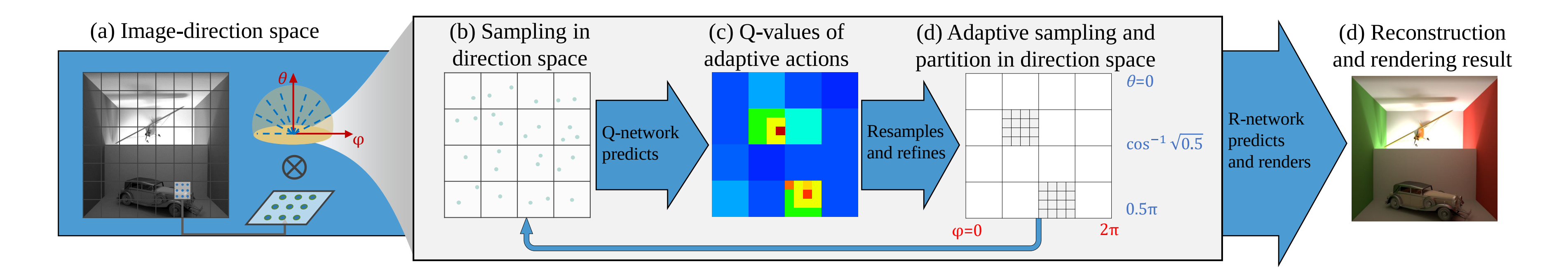


图 1: 光场自适应采样与重建过程 [11]。

2 使用强化学习进行蒙托卡洛应采样与重建

上述经典的自适应采样与重建方法是使用基于先验手动设计的模型在单一场景上进行在线学习，以拟合一个未知的入射光场函数。最近，Huo 等人提出利用跨场景的离线数据来训练深度神经网络 [11]，以更好地进行蒙托卡洛自适应采样与重建。该研究提出训练两个不同的网络，分别称为质量评估网络（Q 网络）和重建网络（R 网络），并使用大量离线数据集来训练两个网络，其中 R 网是一个四维卷积网络，用于将稀疏采样且含有大量噪音的入射辐射场滤波重建为无噪音的密集入射辐射场，Q 网络是则是基于深度强化学习策略训练得到，用于指导获得稀疏样本的采样过程，以提高采样的效率。

整个绘制流水在运行时的流程如图1所示。(a) 首先将图像空间划分为不小的区块。(b) 每个区块内的像素共享的同一个方向空间，使用统一的格子结构离散化方向空间作为初始化，并在每个块中随机生成样本点。(c) 使用训练好的 Q 网络来预测在每个区块内进行“采样动作”所能期望获得的收益。收益是指网络预测出来的进行采样后所能获得的质量提升，在训练 Q 网络时，质量的提升通过比较 R 网络的输出与 Ground Truth 图像的差来获得。“采样动作”包括两种，第一种是加重采样，将区块对应方向空间内的样本数加倍；第二种是细化采样空间，是指将指定的采样空间方向区域划分为粒度更小的格子，以更好地重建区域内的高频细节。在得到预测的价值图后，算法选取可以获得最大收益的动作执行。(d) 展示了将初始方向空间使用细化采样空间的操作后得到的新的采样方向层次结构。(e) 最后使用 R 网络将稀疏的采样点重建为密集的入射辐射场，便可用于路径引导或直接显示等不同的应用。

图2是 R 网络和 Q 网络的训练方法。训练数据集包括 20 个不同的场景，并配以随机生成的视角，并随机修改场景的照明、材质和纹理，一共生成了大约 300 个场景。对于每个配置，生成了由不同采样个数和分辨率的 Ground Truth 光场数据，所包含的每块采样点个数为 $(1, 2, 4, 8, 16)$ 五种，包含的分辨率为 $(16^1, 16^2, 16^3, 16^4, 16^5)$ 五种。这些数据首先被用来训练 R 网络，使其达到收敛并能很好地将稀疏采样重建为密集的无噪光场。收敛后的 R 网络被用来训练 Q 网络，训练同样基于相同的数据集，但是使用了随机强化学习算法来模拟实际采样过程的数据分布情况以防止数据耦合。为了使 Q 网络学习某个采样动作带来的收益，方法将采样动作前后的采样点数据同时送入 R 网络中，以得到两个不同的重建结果，通过将重建结果与 Ground Truth 数据进行比对，可以计算出该采样动作能带来的误差下降幅度，也就是收益。如果某个动作能产生更高的收益，训练后的 Q 网络会引导蒙托卡洛采样过程优先采取该的动作。

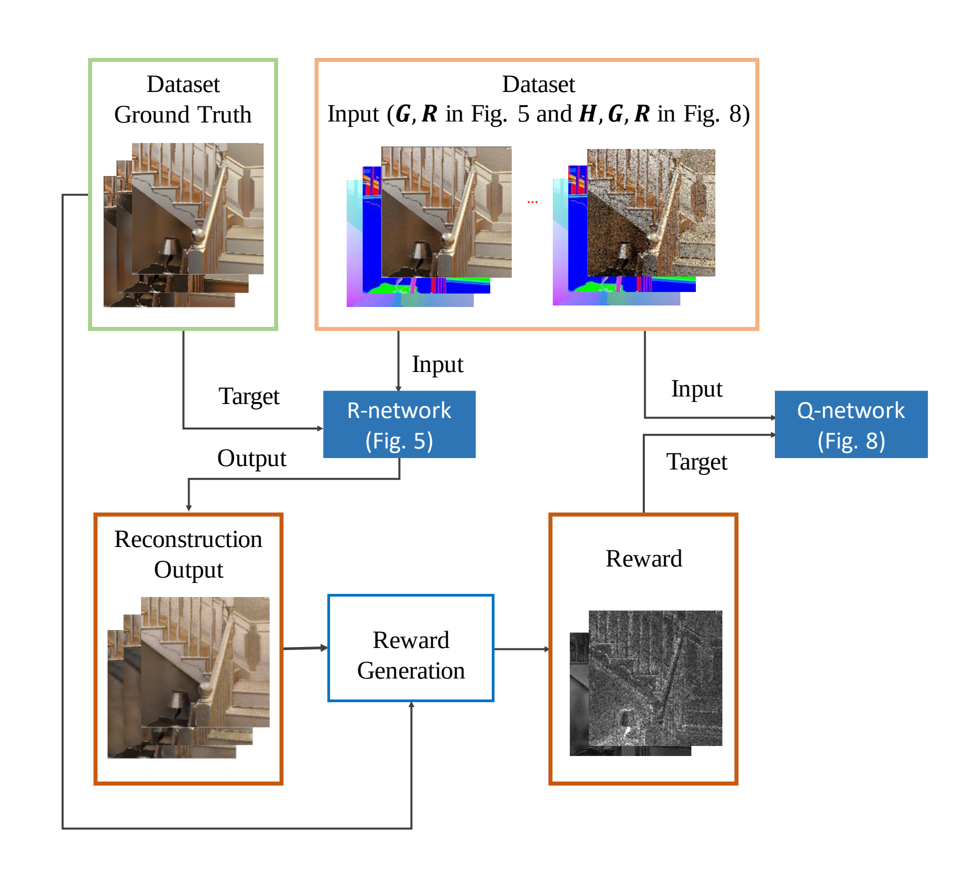


图 2: R 网络和 Q 网络双网络的训练流程 [11]。

3 自适应采样与重建结果

图3展示了基于强化学习的离线学习方法与传统的在线学习方法进行无偏绘制的一个比较。这一场景属于绘制中较难处理的类型，由于半掩的窗子和存在大量遮挡物，路径追踪算法很难生成连接摄像机与光源的有有效光路，场景中充满了大量噪点。普通的路径追踪算法在这一场景中会产生大噪点，而 MEMLT 虽然在开阔区域的生成结果比较平滑，但它的主要问题是几何边界上无法有效收敛，含有大量错误。GMM 和 SD-tree 方法具有相似的工件模式，都是通过在线学习的方法拟合一个入射光场的分布，用于生成 PDF 以引导路径追踪算法。两种方法在开阔区域都可以较好地降噪，但是在含有复杂遮挡的阴影区域，尤其是房间中间的桌子底下，都生成了大量噪点，说明通过简单的共享邻域光线的方法在线重建入射光场，可能会因为不考虑可见性而生成大量错误。新方法与之相比，不仅 Q 网络通过强化学习的策略大量学习了离线数据集的经验数据，可以自适应地预测不同区域的采样效率，同时还通过 R 网络

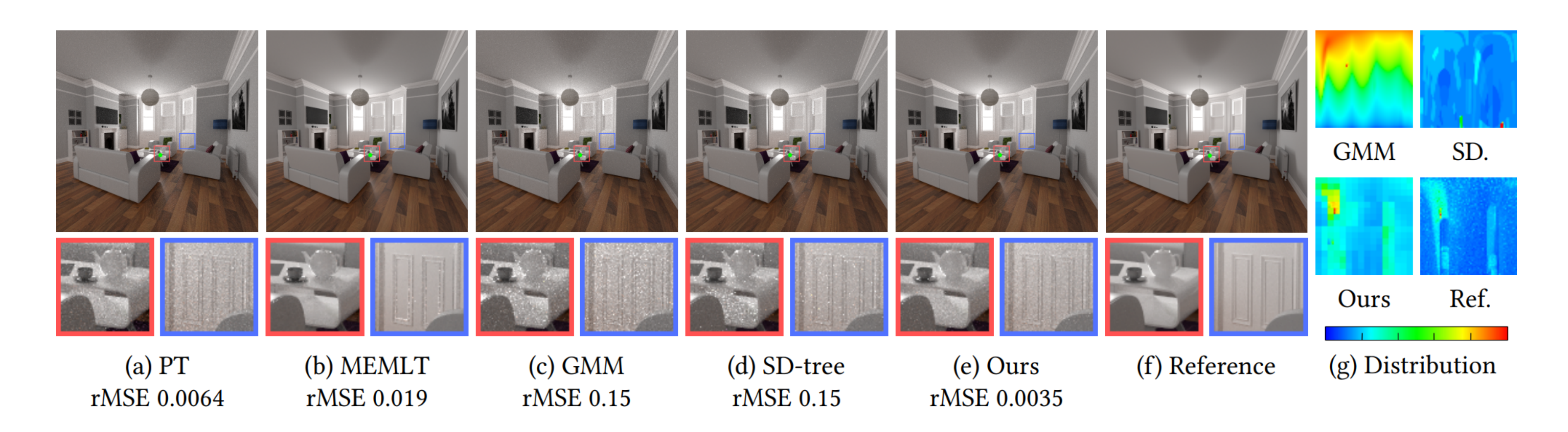


图 3: 基于强化学习的蒙托卡洛自适应光场采样与重建方法与传统无偏绘制算法进行比较。其中图 (a) 是路径追踪算法，图 (b) 是 Metropolis 路径追踪算法，图 (c) 和图 (d) 是基于传统在线学习重建 PDF 函数的路径引导算法 [19, 17]，图 (e) 是文章所提的基于强化学习的蒙托卡洛自适应光场采样与重建进行引导的路径引导算法，图 (f) 是使用 32k SPP 生成的参考图像。图 (g) 是 (c)(d)(e) 三种路径引导算法重建出来的入射光场分布图（绿色点处），通过与参考光场（Ref.）的比较，可看出 GMM 及 SD. 都无法捕捉分布的峰值。[11]

对采样结果进行降噪，有效避免了错误引导及重建结果造成的负面影响，比 GMM，SD-tree 和 MEMLT 的重建误差低了超过一个数量级，比 PT 也低了 2 倍左右。

4 总结与扩展内容

在这项研究中，研究者提出了两种新颖的深度学习网络，R 网络和 Q，用于入射辐射场的自适应采样和重建，这也是首次将强化学习的技术引入蒙托卡洛自适应采样的框架中。除了在无偏绘制和光场重建的应用中取得了最好的结果之外，方法还为图形学领域一些难以生成监督数据的困难问题提供了基于深度学习的解决思路，即通过将复杂问题拆分为简单的原子操作，并通过强化学习的框架来学习一个自适应的渐进优化过程。

与基于物理过程的光能计算相关的推荐扩展内容：

1. 基于深度学习的蒙托卡洛降噪通过离线学习的方法训练一个神经网络降噪器，可以将含噪的蒙托卡洛绘制结果滤波为高质量的光滑输出，大大提升了基于物理的绘制技术的可用性 [12]，常见的研究包括基于 g-buffer 去预测一个滤波核 [1]，使用 GAN 来生成更具真实感的滤波结果 [24]，对路径空间特征进行流形对比学习以强化反射的绘制效果 [3]，使用权重共享的方式快速地预测绘制核以加速重建 [5]，对高维的入射辐射场进行滤波重建以进行

无偏绘制的引导 [11]，等等。

2. 多光绘制框架是路径追踪算法外一个重要的绘制框架，其基本思想是将经过多次折射反射的完整光路照明传输模拟简化为计算来自许多虚拟光源的直接照明的，并提供了统一的数学框架来加速这一运算 [4]，研究包容包括如何高效处理外存中的虚拟点光源和几何数据 [21]，如何使用稀疏矩阵和压缩感知的方法高效积分虚拟点光源 [10]，以及如何处理半透明介质中的虚拟线光源数据 [9]，使用球面高斯虚拟点光源逼近光泽表面的间接反射 [8]，等等。
3. 绘制流水线的自动优化通过优化绘制管线将高质量绘制技术应用与实时的绘制应用中，研究内容包括基于质量和速度的自动优化 [22]，以节能为目标的自动优化 [23, 25]，对地形数据的 LOD 优化 [16]，管线绘制信号自动优化拟合 [15]，等等。

参考文献

[1] S. Bako, T. Vogels, B. McWilliams, M. Meyer, J. Novák, A. Harvill, P. Sen, T. Derose, and F. Rousselle. Kernel-predicting convolutional networks for denoising monte carlo renderings. *ACM Trans. Graph.*, 36(4):97–1, 2017. 4

[2] B. C. Budge, J. C. Anderson, and K. I. Joy. Caustic forecasting: Unbiased estimation of caustic lighting for global illumination. In *Computer Graphics Forum*, volume 27, pages 1963–1970. Wiley Online Library, 2008. 1

[3] I.-Y. CHO, Y. HUO, and S.-E. YOON. Weakly-supervised contrastive learning in path manifold for monte carlo image reconstruction. 2021. 4

[4] C. Dachsbacher, J. Křivánek, M. Hašan, A. Arbree, B. Walter, and J. Novák. Scalable realistic rendering with many-light methods. In *Computer Graphics Forum*, volume 33, pages 88–104. Wiley Online Library, 2014. 5

[5] H. Fan, R. Wang, Y. Huo, and H. Bao. Real-time monte carlo denoising with weight sharing kernel prediction network. In *Computer Graphics Forum*, volume 40, pages 15–27. Wiley Online Library, 2021. 4

[6] S. Herholz, O. Elek, J. Vorba, H. Lensch, and J. Křivánek. Product importance sampling for light transport path guiding. In *Computer Graphics Forum*, volume 35, pages 67–77. Wiley Online Library, 2016. 1

[7] B.-S. Hua and K.-L. Low. Guided path tracing using clustered virtual point lights. In *SIGGRAPH Asia 2015 Posters*, pages 1–1. 2015. 1

[8] Y. Huo, S. Jin, T. Liu, W. Hua, R. Wang, and H. Bao. Spherical gaussian-based lightcuts for glossy interreflections. In *Computer Graphics Forum*, volume 39, pages 192–203. Wiley Online Library, 2020. 5

[9] Y. Huo, R. Wang, T. Hu, W. Hua, and H. Bao. Adaptive matrix column sampling and completion for rendering participating media. *ACM Transactions on Graphics (TOG)*, 35(6):1–11, 2016. 5

[10] Y. Huo, R. Wang, S. Jin, X. Liu, and H. Bao. A matrix sampling-and-recovery approach for many-lights rendering. *ACM Transactions on Graphics (TOG)*, 34(6):1–12, 2015. 5

[11] Y. Huo, R. Wang, R. Zheng, H. Xu, H. Bao, and S.-E. Yoon. Adaptive incident radiance field sampling and reconstruction using deep reinforcement learning. *ACM Transactions on Graphics (TOG)*, 39(1):1–17, 2020. 2, 3, 4, 5

[12] Y. Huo and S.-e. Yoon. A survey on deep learning-based monte carlo denoising. *Computational Visual Media*, 7(2):169–185, 2021. 1, 4

[13] H. W. Jensen. Importance driven path tracing using the photon map. In *Eurographics Workshop on Rendering Techniques*, pages 326–335. Springer, 1995. 1

[14] J. T. Kajiya. The rendering equation. In *Proceedings of the 13th annual conference on Computer graphics and interactive techniques*, pages 143–150, 1986. 1

[15] S. Li, R. Wang, Y. Huo, W. Zheng, W. Hua, and H. Bao. Automatic band-limited approximation of shaders using mean-variance statistics in clamped domain. In *Computer Graphics Forum*, volume 39, pages 181–192. Wiley Online Library, 2020. 5

[16] S. Li, C. Zheng, R. Wang, Y. Huo, W. Zheng, H. Lin, and H. Bao. Multi-resolution terrain rendering using summed-area tables. *Computers & Graphics*, 95:130–140, 2021. 5

[17] T. Müller, M. Gross, and J. Novák. Practical path guiding for efficient light-transport simulation. In *Computer Graphics Forum*, volume 36, pages 91–100. Wiley Online Library, 2017. 1, 4

[18] P. Vévoda, I. Kondapaneni, and J. Křivánek. Bayesian online regression for adaptive direct illumination sampling. *ACM Transactions on Graphics (TOG)*, 37(4):1–12, 2018. 1

[19] J. Vorba, O. Karlík, M. Šik, T. Ritschel, and J. Křivánek. On-line learning of parametric mixture models for light transport simulation. *ACM Transactions on Graphics (TOG)*, 33(4):1–11, 2014. 1, 4

[20] J. Vorba and J. Křivánek. Adjoint-driven russian roulette and splitting in light transport simulation. *ACM Transactions on Graphics (TOG)*, 35(4):1–11, 2016. 1

[21] R. Wang, Y. Huo, Y. Yuan, K. Zhou, W. Hua, and H. Bao. Gpu-based out-of-core many-lights rendering. *ACM Transactions on Graphics (TOG)*, 32(6):1–10, 2013. 5

[22] R. Wang, X. Yang, Y. Yuan, W. Chen, K. Bala, and H. Bao. Automatic shader simplification using surface signal approximation. *ACM Transactions on Graphics (TOG)*, 33(6):1–11, 2014. 5

[23] R. Wang, B. Yu, J. Marco, T. Hu, D. Gutierrez, and H. Bao. Real-time rendering on a power budget. *ACM Transactions on Graphics (TOG)*, 35(4):1–11, 2016. 5

[24] B. Xu, J. Zhang, R. Wang, K. Xu, Y.-L. Yang, C. Li, and R. Tang. Adversarial monte carlo denoising with conditioned auxiliary feature modulation. *ACM Trans. Graph.*, 38(6):224–1, 2019. 4

[25] Y. Zhang, R. Wang, Y. Huo, W. Hua, and H. Bao. Powernet: Learning-based real-time power-budget rendering. *IEEE Transactions on Visualization and Computer Graphics*, 2021. 5